EDUC632A: FUNDAMENTALS OF TEST CONSTRUCTION

Stephen G. Sireci, Ph.D.
156 Hills South
sireci@acad.umass.edu
http://www-unix.oit.umass.edu/~sireci
(413)545-0564 (voice)

Office Hours:
Monday: 1:30—3:00, Friday: 1:30—3:00
Other times by appointment

Course Syllabus and Schedule for Fall 2014

Course Objectives: This course will provide information on how to build and evaluate educational tests and how to effectively and appropriately interpret test results. Students will learn about the advantages and disadvantages of different assessment formats such as selected response items, performance assessments, and computer-based testing. Specifically, students will learn how to:

- describe fundamental aspects of test quality such as reliability and validity
- operationally define testing purposes
- develop a variety of item formats including multiple-choice and constructed-response items
- develop answer keys and scoring rubrics for different item formats
- evaluate tests and items using statistical and qualitative methods
- incorporate meaning into test score scales using both norm-referenced and criterion-referenced procedures
- use standard setting techniques to set “passing scores” and other performance standards on tests
- develop appropriate documentation to properly communicate the quality of an assessment
- understand the utility of educational assessments within the broader context of educational policy and decision making

The common theme uniting these knowledge and skill areas is the promotion of equity and fairness in educational testing. In addition, the course stresses the role of educational testing in improving student learning. Students will learn how to build quality tests aimed towards promoting valid score interpretation, and will learn how to evaluate the use of a specific test for a specific purpose. Measuring psychological phenomena such as what a student “knows and is able to do” is a complex endeavor. Test construction is both art and science; both aspects will be stressed in this course. Upon successful completion of this course, students will know how to (a) develop tests, (b) choose among already existing tests for a specific purpose, (c) use the results of standardized tests to help make decisions about students and educational systems, and (d) identify flaws in educational assessments.

Some specific topics covered in the course are:
- Purposes of Educational Tests
- Standards for Teacher Competence in Educational Assessment
- Standards for Educational and Psychological Testing
- Fundamental Elements of Test Quality (e.g., reliability, validity)
- Developing Multiple-Choice Items
- Developing and Scoring Constructed-Response Items
- Developing Portfolio Assessments
- Item Analysis
Specific topics covered (continued):
  o Evaluating the Validity of Score-Based Inferences
  o Standard Setting
  o Innovative Item Formats and Computer-Based Testing
  o Test Accommodations for individuals with disabilities and for English learners
  o Sensitivity Review
  o Ethical Issues in Test Construction, Selection, Administration, and Interpretation

**Course Requirements**

A. **Attendance and Participation:** Students expecting to receive course credit will need to attend all (or nearly all) classes, work their way through the suggested readings, and complete several assignments. In addition, students are expected to actively participate in class.

B. **Assignments:** In addition to weekly homework assignments, there are two major assignments for the course:

  1) **Test Development:** Each student is required to develop an educational achievement test. This semester, we will be developing math and reading tests for adult basic education students in Massachusetts. Throughout the semester, students will progressively work on the design and development of their test. This process will involve writing items that may eventually be field-tested on adult education students and end up on the Massachusetts Adult Proficiency Tests.

  2) **Technical Manual Development:** Each student is required to develop a technical manual that describes the development process for their test and provides important psychometric data for their test. Instructions for developing this manual will be provided in class. Chapter 7 of the *Standards for Educational and Psychological Testing* (AERA, APA, & NCME, 2014) should be used to guide development of the manual.

**Grading:** Students’ final grades are determined by their attendance and participation in class, and by their performance on their weekly assignments, draft test, final test, and technical manual. Late assignments will be reduced by one-letter grade for each day late (e.g., a maximum grade of "C" will be given to an exceptional draft test submitted two days late). Unforeseen emergencies, as determined by the professor, are exceptions to this policy. The table below illustrates the weighting used in calculating grades.

<table>
<thead>
<tr>
<th>Activity</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attendance/Participation</td>
<td>.15</td>
</tr>
<tr>
<td>Weekly assignments</td>
<td>.30</td>
</tr>
<tr>
<td>First Test Draft</td>
<td>.05</td>
</tr>
<tr>
<td>Final Test Form</td>
<td>.25</td>
</tr>
<tr>
<td>Technical Manual</td>
<td>.25</td>
</tr>
</tbody>
</table>

Attendance/participation and all assignments are graded on a 0-100 scale. Final grades of 94-100 receive an A, 90-93 receive an A-, 87-89 receive a B+, 81-86 receive a B, 79-80 receive a B-, 77-78 receive a C+, 70-76 receive a C, and below 70 receive an F.
Suggested Textbook

This is a terrific book. If this is your first course in educational measurement, you should get it. If you have other books that cover test construction, you may not need it because I will provide numerous handouts on all topics covered in the course. There are 10th and 11th editions of this book, but I am using the 9th edition because it contains essentially the same material and is about 25% of the cost of the 11th edition. Feel free to use the 9th or a later edition. You can also use Gallagher’s (1997) text *Classroom Assessment for Teachers*, but it is a bit dated.

Recommended Text

Required Readings
Each week you will have one or more reading assignments. Most of these readings are listed in the bibliography that appears next. I will distribute or post copies of all reading assignments. Articles that are likely to appear as reading assignments are denoted with an asterisk (*). The bibliography is stratified by topic area to facilitate finding publications in specific areas of interest.

Fundamentals of Test Construction Bibliography

**Item Development/Formats**


Validity


**Standard Setting**


**Computer-Based Testing**


**Testing Special Populations**


**Performance Assessment**


**Cognitive/Principled (Evidence-Centered) Assessment Design**


**Portfolio Assessment**


**Miscellaneous**


**Academic Honesty Statement**

The integrity of the academic enterprise of any institution of higher education requires honesty in scholarship and research; thus academic honesty is required of all students at the UMass. Academic dishonesty is prohibited in all programs of the University. Academic dishonesty includes but is not limited to: cheating, fabrication, plagiarism, and facilitating dishonesty. Appropriate sanctions may be imposed on any student who has committed an act of academic dishonesty. Any person who has reason to believe a student has committed academic dishonesty should bring the information to my attention. Instances of academic dishonesty not related to this course should be brought to the attention of the appropriate department Chair. Students are expected to be familiar with this policy and the commonly accepted standards of academic integrity, and so ignorance of such standards is not normally sufficient evidence of lack of intent.

**Plagiarism policy:** Direct copying of someone else’s work is not allowed. Printing out someone else’s computer output, and handing it in as your own work, is also not allowed. Passing off someone else’s work as your own will result in failing this course. Please see me if you have questions about this policy, or if you have trouble completing any assignments.

**Accommodation policy:**
I strive to provide an equal educational opportunity for all students. If you have a physical, psychological, or learning disability, you may be eligible for academic accommodations to help you succeed in this course. If you have a documented disability that requires an accommodation, please notify me as soon as possible, but no later than the third class, so that we may make appropriate arrangements to provide any needed accommodations for class or assignments.
**TENTATIVE CLASS SCHEDULE FOR FALL 2014**

Listed below are the topics that will be covered as well as a list of suggested readings. The dates listed for each topic are tentative.

<table>
<thead>
<tr>
<th>Class</th>
<th>Topics</th>
<th>Readings</th>
</tr>
</thead>
<tbody>
<tr>
<td>9/2</td>
<td>Purposes of Educational Tests</td>
<td>Brookhart (2011); Cizek (2001)</td>
</tr>
<tr>
<td></td>
<td>Standards for Teacher Competence in Assessment</td>
<td>Text Chs. 1-2, &amp; Appendix D</td>
</tr>
<tr>
<td></td>
<td>Norm- and Criterion-referenced testing</td>
<td></td>
</tr>
<tr>
<td>9/9</td>
<td>Reliability and Validity Fundamentals</td>
<td>Sireci (2005), Text Chs., 3-5</td>
</tr>
<tr>
<td>9/16</td>
<td>Defining Test Content and Developing Test Specifications</td>
<td>Text Ch. 6, Handouts</td>
</tr>
<tr>
<td><em>9/23</em></td>
<td>Assessment Format Options</td>
<td>Text Ch. 8</td>
</tr>
<tr>
<td></td>
<td>Writing Multiple-Choice (MC) Items</td>
<td>Haladyna &amp; Downing (1989)</td>
</tr>
<tr>
<td>9/30</td>
<td>Writing MC Items (continued)</td>
<td>Text Ch. 7</td>
</tr>
<tr>
<td></td>
<td>Writing Other Objectively Scored Items</td>
<td>Haladyna (1992)</td>
</tr>
<tr>
<td>10/7</td>
<td>Developing Performance Assessments</td>
<td>Text Chs. 10-11; Dunbar et al. (1991); Linn &amp; Burton (1994)</td>
</tr>
<tr>
<td>10/14</td>
<td><strong>No Class—Monday Schedule @ UMASS</strong></td>
<td></td>
</tr>
<tr>
<td>10/21</td>
<td>Scoring Performance Assessments</td>
<td>Ch. 11, Handouts</td>
</tr>
<tr>
<td></td>
<td>Sensitivity Review</td>
<td>Sireci &amp; Mullane (1994)</td>
</tr>
<tr>
<td>11/4</td>
<td>Field Testing and Item Analysis</td>
<td>Text Ch. 14, Handouts; Wainer (1989)</td>
</tr>
<tr>
<td><em>11/12</em></td>
<td><strong>NOTE: This class is a WEDNESDAY</strong></td>
<td>Text Ch. 19</td>
</tr>
<tr>
<td></td>
<td>Incorporating Meaning Into the Test Score Scale</td>
<td>Cizek (1996)</td>
</tr>
<tr>
<td></td>
<td>Setting Standards on Educational Tests</td>
<td></td>
</tr>
<tr>
<td>11/18</td>
<td>Understanding Test Results</td>
<td>AERA, APA, &amp; NCME (2014)</td>
</tr>
<tr>
<td></td>
<td>Developing a Technical Manual</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Wainer (1993); Zenisky &amp; Sireci (2002)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Thompson &amp; Thurlow (2002); Text Ch. 12; Huff &amp; Goodman (2007)</td>
</tr>
<tr>
<td>12/16</td>
<td><strong>Final Test Form and Technical Manual Due (no class)</strong></td>
<td></td>
</tr>
</tbody>
</table>

*Indicates Online Class*

At this point, we anticipate 4 online classes, with students having an option to participate remotely in a 5th class via Fuze. Details to be discussed during the first class.